HUMAN ACTIVITY

LOCALIZATION AND RECOGNITION

BASED ON RADAR SENSORS FOR SMART HOMES

A THESIS

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENT

FOR THE DEGREE

MASTER OF SCIENCE

BY

SHANGYUE ZHU

DR. SHAOEN WU- ADVISOR

BALL STATE UNIVERISTY

MUNCIE, INDIANA

DECEMBER 2017

Contents

1 INTRODUCTION

With the development of Internet of Things and technology on smart home, realization of health-care system for medical industry is desired as well as a security system [24]. There are many potential damages to school-age children and elderly people. For instance, the potential crash that will occur when children jog or do more intensive motion in the home [14]. And elderly people have the potential to fall at any moment, when they try to stand up or sit down [24]. Hence, utilizing the smart sensors to recognize the human activities has become an active area of research. In this area, smart home system has an advance development in computing, ambient intelligence and miniaturization of technology to provide a quality life.

In smart systems, one critical capability required is recognizing residents activities in their daily lives. Some smart systems are supported by camera videos, and more smart systems recognize the users activities depending on the wearable sensor data. However, the camera video has potential privacy invasions [32]. Because they often take off or forget to wear sensors, information on elderly people and school-age children is always unreliable when collected. Thus, it is of ultimate interest to design a passive and non-invasive IoT system for an active smart environment.

In this thesis, we propose a smart home system including two main functions: tracking a user's location and recognizing the activity. For tracking users' location, we propose a IoT solution, Distance based Localization and Tracking (DiLT), which collects data by using commodity off-the-shelf ultrasonic sensor for minimal invasion to localize and track a user in an environment. DiLT uses *distance* as special "signal" and adopts high-degree signal processing techniques to reveal detail dynamics embedded in the sensed

data. The algorithm has the following highlights: **First,** to address the challenge of noise in ultrasonic measurement, we design a robust mechanical beamforming ultrasonic system that expands the ultrasonic sensing range and capability of conventional ultrasonic sensors. The mechanical beamforming ultrasonic system can collect direction-aware data and suppress noise in data measurement. **Second,** this work designs a data processing algorithm to remove exceptional and noisy data from the collected data, and transform the multiple sampled data sets into a single data set. **Third,** we design a *contrastive divergence learning* algorithm to localize a user based on the sensed distance data with high accuracy. **Forth**, this thesis proposes a *Binary Backoff* (BNB) algorithm to track down the change of user location.

For recognizing the activity, we propose a solution of indoor Human Activity Recognition based on Ambient Radar sensors, *HARAR*. This work considers four types of indoor activities: *sit-to-stand, stand-to-sit, walking* and *jogging*. The algorithm has the following highlights: **1.** it proposes a chain of signal processing algorithms to: 1) remove the exceptional measurement and interpolate the replacement, 2) filter out the static background reflections to keep only the motion reflections, and 3) transform the signal data into relative location changes with rich features. **2.** it designs: 1) an algorithm to separate the sequence of activity mixture into individual activities, 2) a lowpass filtering algorithm to remove the unwanted noisy components in the data for accuracy, and 3) a motion intensity based classification method to separate *sit-to-stand* and *stand-to-sit* from *walking* and *jogging*. **3.** it uses miniature radar to emit signals at 16 pulse per second while the measurement of the reflected signal occurs at a very high frequency of 128 KHz to capture very fine dynamics of activities. **4.** it invents two features: slope gradient and relative

velocity to recognize each activity finely with the k-means machine learning algorithm.

In the remaining of this thesis, the brief overview of activity recognition and related work are presented in Section 2. Section 3, introduces the basic backgournd including fundamentals of testing device and type of collected data. Section 4 next discusses the detail system design of DiLT, including the physical design, data processing, localization and tracking algorithms. Section 5 discusses the detail system design of *HARAR*, including the data preprocessing, coarse activity classification and fine activity recognition. In Section 6, extensive evaluation results have been presented. Section 7 concludes this work.

2 RELATED WORK

2.1 Activity Recognition

Human activity recognition is a kernel construction block behind smart home applications. It takes the raw sensor reading as inputs to predicts a residents motion [29, 34]. Many previous researches are equipped with various sensors, including accelerometers [30], Gyroscope [12], light sensors [28], temperature sensors [8], video [15, 22], etc. There sensors become a rich data source to record various aspects of residents life. For instance, wearable sensors are designed to be worn on human body in daily activities. They record residents physiological states, such as step changes, moving directions, speed, etc [16, 30, 34]. In the [16], the authors described utilizing a body worn wireless accelerometer to be used in the real-life application of patient monitoring. Their algorithm collect data from a single, waist-mounted triaxle accelerometer to classify gait events into

multiple daily living activities.

Moreover, many researchers have explored activity recognition base on multiple algorithms analysis [9, 13, 37]. In [37], the authors investigate daily living analysis from visual data gathered from wearable cameras. In particular, they use multitask clustering algorithm derived motion features from complex images data. The authors of [13] focus on discussing prediction algorithms to bring about next event recognition. Their Episode Discovery helps in finding the frequency of occurrence of particular events.

2.2 Radar Sensor

Radar sensor has been widely used as a sensing modality for interactive systems and applications, because they do not depend on lighting, noise or atmospheric conditions [26, 35]. Google ATAP team has designed an mm-wave radar system Soli [27] based on 60GHz signals to capture subtle motions in gestures. Soli receives reflected signal and use signal processing techniques to extract features and employs machine learning to recognize different gestures with classification algorithm. Moreover, in [38], the authors investigate the use of Doppler radar sensor for occupancy monitoring. They extract different levels of activity, which is detected by post-processing sensor signals, to monitoring the situation of room occupancy.

2.3 Ultrasonic Sensing

The fundamental principle of ultrasonic is similar to radar sensing on radio waves. The ultrasonic wave has a frequency higher than the frequency of the sound wave by more than 20 kHz. Ultrasonic sensors consist of an ultrasonic transmitter and a receiver. In

working, the ultrasonic signal is emitted from the ultrasonic transmitter. When the signal hits an object, the signal is reflected and received by the ultrasonic receiver [5]. The signal is then delivered to a micro-controller for further processing to calculate the distance to the object.

2.4 Ultrasonic Locaization

Previous efforts have been attempted to use ultrasonic for localization. Two of such efforts most related to our work are the solutions proposed by Nishida *et al* [31, 32]. In the first work [32], they have multiple ultrasonic transmitters and receivers installed in a celling. This ultrasonic radar system detects human motion at a relatively high vertical position to recognize the location of residents at a home. They additionally use the reflected sound pressure to visualize the position of human in detection area. In their other work [31] extending their first work, Nishida *et al* propose a three-dimensional ultrasonic localization and tagging system. They use multiple ultrasonic receivers embedded in both wall and ceiling to calculate a three-distance measurement from a tri-lateration detection area. Moreover, Angelis *et al* [10] also investigate three-dimensional positioning based on ultrasonic sensors. They vary the sampling frequency by selecting different transmitters and receivers to track an object dynamically. Meanwhile, they also explore the miniaturization of the ultrasonic transmitters and developed a systematic method for defining and tagging activities with high accuracy. In addition, ultrasonic sensing has been also used in smart systems other than localization [1, 6, 11, 17, 25, 33, 36, 39].

Many researchers have explored localization based on ultrasonic context-aware computing. In [21], the authors investigate indoor localization with RF systems and ultra-

sonic. The indoor positioning system is accessed by multiple ultrasonic emitters. The location is estimated with different ultrasonic signal reflections through marked location points. The authors of [4] develop a portable device to synchronize corresponding ultrasonic location code in a period. The device is configured to receive the timing synchronized information and to transmit a location code based on the received timing synchronization. Google ATAP team has designed an mm-wave radar system Soli [27] based on 60 GHz signals to capture subtle motions in gestures. *Soli* receives reflected signal and use signal processing techniques to extract features and employs machine learning to recognize different gestures with classification algorithms.

2.5 Beamforming

Beamforming has been widely used as a flexible signal processing technique usually in sensor arrays for directional wireless communications [2]. Medical systems have used beamforming in ultrasound imaging. For instance, a system has been proposed to use the synthetic aperture sequential beamforming (SASB) technique for clinical patient scanning [18]. This system can extract the cancer features from limited image information based on SASB with high quality image.

3 BACKGROUND

In this section, we propose to introduce the fundamentals of radar sensor, the test used in this research, and the type of collected data.

3.1 Radar Fundamentals

Radar sensors transmit radio waves, which can be detected by the radar receiver to illuminate the target in its detection range. In working, the radar signal hits an object, the signal is reflected and received by receiver terminal [24]. The signal is then delivered to a micro-controller for further processing to analyze the amplitude and other parameters for the object.

3.2 Device Selection

In order to implement observed experimental results, we select two different devices for the distinct experiment parts. For tracking users location, we selected the ultrasonic in our system is an HC-SR04 ultrasonic sensor that can detect a distance of 5 to 400 cm (or 2 to 156 in)[1]. The servomotors are Micro servo that have an active range of 0 to 180 degree.

For activity recognition, this sensing platform is based on Walabot[2] that has a size of 72 mm $\times$ 140 mm. Walabot supports multiple antenna pairs to sense a target area and each pair consists of two directional antennas working on different frequency ranges. FCC regulates the wireless operates over 3.3-10.3 GHz range. The average transmission power of both models is about -16 dBm. In this research, in order to simulate an ordinary home circumstance, we configure antennas to work on a medium frequency 7.85GHz. The radar field of the view is approximately 60 degrees horizontally and 15 degrees vertically.

[1]Although the sensor specification indicates the range 2 - 400 cm, our actual validation shows the reliable range is 5 - 400 cm

[2]https://walabot.com/community

3.3 K-means Clustering Algorithm

The k-means algorithm is the process of partitioning a set of data objects (or observations) into subsets. Each subset is called a cluster, such that objects within a cluster are similar to one another, yet dissimilar to objects in other clusters. It has a loose relationship to the k-nearest neighbor classifier, a popular machine learning technique for classification that is often confused with k-means because of the k in the name [19]. Cluster analysis has been widely used in many applications such as image pattern recognition and Web search. Partitioning methods, which are distance-based and use mean to represent cluster center, are effective for small-size to medium-size data sets. In this thesis, we utilize the K-means Clustering Algorithm focus on analysis of the motion features such as velocity in activity recognition part.

4 LOCALIZATION SYSTEM DESIGN

The core idea of DiLT consists of three main parts: (1) collecting static environment background data, (2) gathering instantaneous scenery data, and (3) analyzing the differentiation between the scenery and the background data sets to locate and track an occupant. This section presents the various research challenges that DiLT has to address, the methodologies that DiLT employs and its system components.

4.1 Physical Design

The capability of ultrasonic sensors is so constrained that they can only detect the distance of an object directly in the front. As a result, an ultrasonic sensor can only work

on a very limited angle or space. The first research challenge is *to design an ultrasonic system to cover the entire space of a room.* We design a mechanical ultrasonic beamforming system by using an off-the-shelf commodity ultrasonic sensor, two servomotors and an Arduino UNO board, as shown in Figure 1.

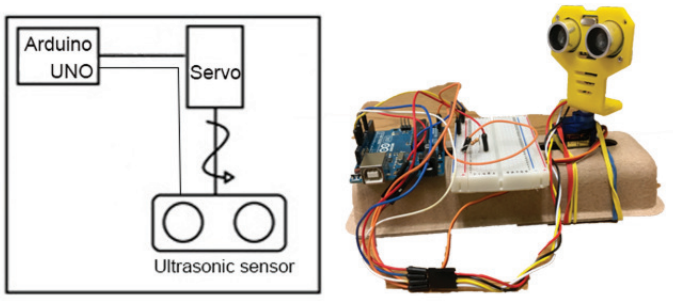

Figure 1: System Architecture of DiLT

The result system concept diagram is plotted in Figure 2. This particular design allows each measurement to be collected as a pair of semicircle coordinates (*motorangle*, *distance*), which is the data format to be processed.

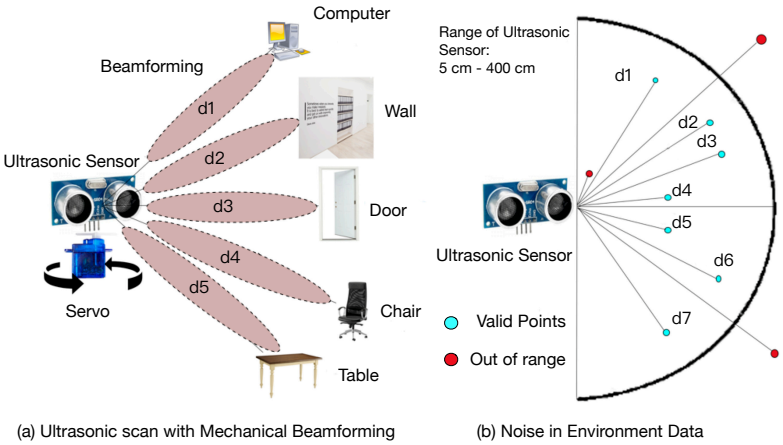

Figure 2: Ultrasonic Beamforming Concept

4.2 Data Collection and Formatting

The data collection includes gathering both static environment background data and dynamic scenery data. In particular, the motors repeat the rotation from 0° to 180° and then back to 0° for several loops. As a result, each data entry is presented by a pair of metrics (*motorangle*, *distance*). The data collected in one task cycle of several repeated rotation loops is formatted as a distance matrix $X \in R^{y \times n}$, where y refers to the number of scanned spots (or angles) in a single rotation loop, and n refers the number of scan rotation loops in a task cycle. $x_{i,j}$ is therefore the distance measured for the ith angel in the jth rotation loop.

$$X = \begin{bmatrix} x_{00} & x_{01} & \cdots & x_{0n-1} \\ x_{10} & x_{11} & \cdots & x_{1n-1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{y0} & x_{y1} & \cdots & x_{yn-1} \end{bmatrix} \tag{1}$$

4.3 Data Processing

After the raw sensor data are obtained, a serial of data processing has to be performed before they can be mined for observations.

4.3.1 Exceptional Data Removal

It is reasonable to expect that the collected ultrasonic data is noisy because of two causes: (1) the coarse resolution and accuracy of an off-the-shelf ultrasonic sensor, and (2) the irregular surfaces of static objects in an environment. A research challenge is, therefore, to clean the data by removing the noisy data.

The first part is to remove the noisy data out of measurement range resulted from the

accuracy of the sensor and device. For removed noisy data, we replace it with a fabricated data so that this location will not loss its information in the data set. Assuming that two neighbor scanning positions are similar in information, DiLT simply uses the data of its previous scan as the fabricated data for this current position that has noisy data removed. Mathematically, referring to the data format matrix Formula (1), if $x_{i,j} \geq 400$ or $x_{i,j} \leq 5$, then $x_{i,j} = x_{i-1,j}$. After this data preprocess, all data should fall into the valid measurement range of the sensor.

The second part of the data cleaning is to remove the outlier data caused by irregular object surfaces. Our system adopts 50 cm as the threshold to claim a spike outlier because such a distance variation is large enough for an object with a width of only 14 cm to be considered abnormal. With this threshold, the outlier removal is performed as: if $x_{i,j} - x_{i-1,j} \geq 50$ and $x_{i,j} - x_{i+1,j} \geq 50$, then $x_{i,j} = (x_{i-1,j} + x_{i+1,j})/2$. In this filtering algorithm, the removed outlier data is replaced with the mean of its previous and successive scanned data, namely the irregular surface spot is emulated by a fabricated smooth transition spot.

4.3.2 Data Transformation

After the exceptional data are moved, the measured data in one task cycle are still in the format of a matrix shown in Equation (1). The i-th row $(x_{i,0}, x_{i,1}, ..., x_{i,n-1})$ contains the data measured for the same spot i (or angle) in all n repeated scan loops.

For the transformation, rather than using the mean of the n loops to for each scan spot that takes in inaccurate measurements, our system performs a "Mode" operation on each row of the matrix. The "*Mode*" operation, by definition, outputs the number

which appears mostly often in a set of numbers. For example, Mode({2, 3, 2, 6, 2, 5, 2, 3}) gives 2 because it appears most often in the set. The transformation with "Mode" operation basically uses the most "likely" reliable measurements in these n repeats as the distance of a scan spot. Denote m_i the output of the Mode operation on the i-th row of the matrix. We have the transformation as:

$$m_i = Mode(x_{i0}, x_{i1}, ..., x_{in-1})$$

To further accommodate the likely *minor* variations in repeated loops, in the transmission, the number at ones of x_{ij} is deliberately omitted; namely 57 is for example considered as 50. The transformation output is a vector M that represents the final collected data in a task cycle scanning y spots or angles. The entire transformation is illustrated as below.

$$\begin{aligned} M &= Tr(X) = Tr(\begin{bmatrix} x_{00} & x_{01} & \cdots & x_{0n-1} \\ x_{10} & x_{11} & \cdots & x_{1n-1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{y0} & x_{y1} & \cdots & x_{yn-1} \end{bmatrix}) \\ &= \begin{bmatrix} Mode(x_{00} & x_{01} & \cdots & x_{0n-1}) \\ Mode(x_{10} & x_{11} & \cdots & x_{1n-1}) \\ \vdots & \vdots & \vdots & \vdots \\ Mode(x_{y0} & x_{y1} & \cdots & x_{yn-1}) \end{bmatrix} = \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_y \end{bmatrix} \end{aligned}$$

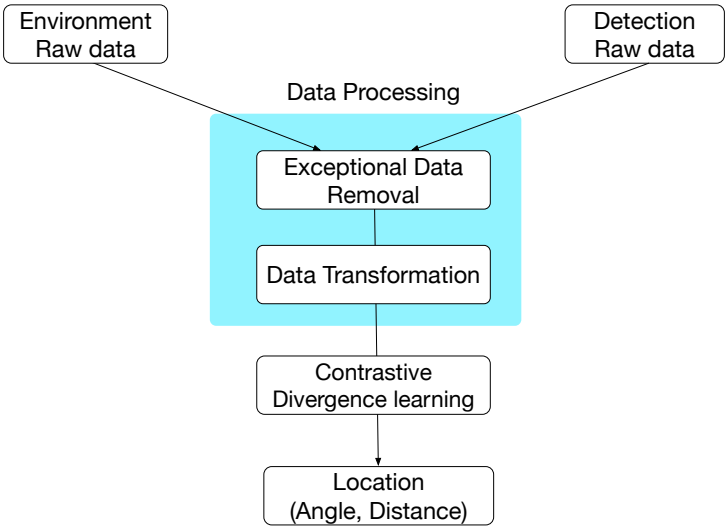

Figure 3: DiLT Localization

4.4 User Localization and Tracking

Our DiLT system localizes and tracks a user in a two-step procedure: (1) extracting the distance data incurred by user activities from the measured distance mixture of the user and the static environment, and (2) mining the extracted distance data of user activities to determine the location of a user and track the location change. Before the localization and tracking is performed, the data vector M_{env} of static environment background supposes to have been obtained through the data processing discussed above in Section 4.3.

4.4.1 Localization

After DiLT performs a scanning task on the environment at a certain moment, the data processing will generate a vector M_{mix} that contains the mixture information of the user and the environment background. Denote M_{user} the user data. M_{mix} is a combination of M_{env} and M_{user}. With the availability of M_{env} and M_{mix}, M_{user} is extracted by a *contrastive divergence learning* (CDL) algorithm [7] as illustrated on Figure 3.

The CDL algorithm can detect one divergent data of a scanned spot in M_{mix} that

is different from the static background data M_{env} and extracts this divergent data as a target data. After all divergent data are obtained, based on the consecutiveness of the divergent data, the approximate location can be inferred: *the user locates across the angles that correspond to the consecutive divergent data.*

4.4.2 Location Tracking

One scan task can determine the location of a user with the localization algorithm. Multiple such scan tasks can therefore generate a *streak* of locations of a certain user, namely tracking the user location changes. In our proposed DiLT system, to track the motion of a user, after the user is localized, a set of successive scan tasks will be performed on the adjacent spots of a user localized. It is important to determine how many of the adjacent spots should be scanned: on one side, a too small adjacent region likely misses the target, and on the other side, a too large region hurts the system efficiency.

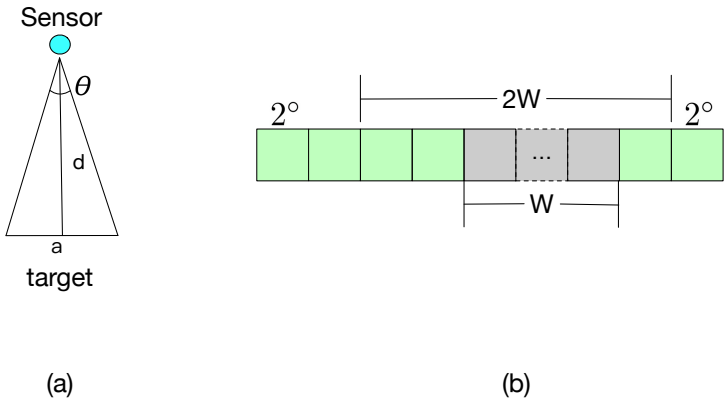

Figure 4: Location Tracking

To determine the adjacent region size, the system first calculates how many scanning spots that the localized user expects to occupy. Referring to Figure 4(a), denotes a the target width (aperture), d the distance between the target and the sensor that can be obtained from the measurement. Then, the angles θ occupied by the target can be

determined as in Equation (2). In our system, because each scan rotates 2°, the number of spots W occupied by the target is determined by Equation (3).

$$\theta = \arctan(\frac{a/2}{d}) \times 2 \quad (2)$$

$$W = \frac{\theta}{2} = \arctan(\frac{a/2}{d}) \quad (3)$$

After the occupying size W is obtained, the system uses a *binary back-off* (BNB) algorithm to decide the adjacent region size as shown in Figure 4(b). Specifically, the system uses $W^1 = 2W$ as the initial adjacent scan region size to identify the new location of the target in the next scan task. If the target is inside the adjacent region, the new user location can be identified and the new adjacent region will be iteratively based on the new location in the next tracking task. Otherwise, the adjacent region size is too small and the size will be doubled to $W^2 = 2 \times W^1$ to localize the moving target, and this BNB will be repeated until the location of the user is identified.

5 ACTIVITY RECOGNITION SYSTEM DESIGN

To enable smart environments with commodity ambient sensors, we propose a solution: Human Activities Recognition Based on Ambient Radar (*HARAR*). *HARAR* repeats emitting a 7.8 GHz wireless radio signal about every 0.7 second through a sending antenna. Meanwhile, it actively measures the signal power reflected by human body parts with an array of receiving antennas. It then employs signal processing and machine learning methods to accurately recognize human activities based on the patterns of reflected signal power. We present the detailed system design of *HARAR* in this section, including

(a) data collection and (b)human activity recognition.

5.1 Sensing Platform, System Model and Data Collection

HARAR uses a continuous-wave radar sensing platform to collect human activity data by following the radar principles [24]. In particular, this sensing platform is based on Walabot[3] that has a size of 72 mm × 140 mm. Walabot supports multiple antenna pairs to sense a target area and each pair consists of two directional antennas working on different frequency ranges. FCC regulates the wireless operates over 3.3-10.3 GHz range. The average transmission power of both models is about -16 dBm. In this research, in order to simulate an ordinary home circumstance, we configure antennas to work on a medium frequency 7.85GHz. The radar field of the view is approximately 60 degrees horizontally and 15 degrees vertically.

Our radar sensing platform emits probing pulse signals $x(t)$ at a pulse repetition frequency (PRF) of 16 Hz, but within each pulse repetition interval (PRI), the receiver antenna samples the received signal $y(t)$ at a very high frequency of 8 KHz. Considering human activities usually stay on lower frequencies, such a radar sensing is capable enough to catch activity dynamics. Although Walabot provides APIs to get preprocessed data at coarse resolutions, it does provide a mechanism to extract raw received signal amplitudes at its native analog-to-digit (ADC) rate, which offers the highest resolution of data. Our system exactly uses this raw data option to obtain the highest possible resolution data of signal variation dynamics during human activities.

When a human object is within the detection area, body parts can be modeled as

[3]https://walabot.com/community

a collection of reflective points, as shown on Figure 5. The emitted signal $x(t)$ arrives at and is then modulated by body parts independently. As a result, the radar signal signature $y(t)$ at a receiver antenna is a mixture of those modulated signals. The posture changes of body parts in various activities are expected to result in different patterns in the radar signal signatures. Therefore, by analyzing the radar signal signature patterns, activities are expected to be recognized.

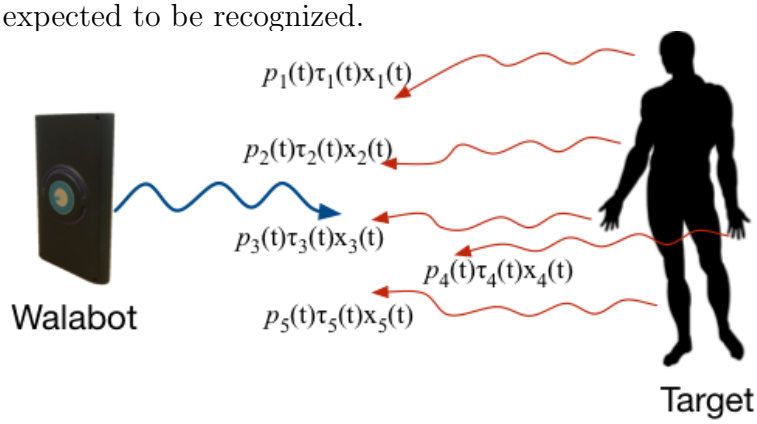

Figure 5: Radar Sensing

We model the radar frequency response of human activities as a superposition of responses from a collection of N various discrete body scattering points, which can be formulated as:

$$y(t) = \sum_{0}^{N-1} \rho_i(t) r_i(t) x(t) \tag{4}$$

where $\rho_i(t)$ is the complex reflectivity parameter of the body reflective point i, $r_i(t)$ presents the corresponding round-trip channel response between the radar sensor and the body reflective point i.

It should be noted that the received signal $y(t)$ contains not only reflections from body parts, but also those from environment background. We define background reflections as

"noise" $n(t)$ to the activity signals. Then $y(t)$ follows:

$$y(t) = \sum_{0}^{N-1} \rho_i(t) r_i(t) x(t) + n(t) \tag{5}$$

As the received signal signature $y(t)$ carries the information of body activities, we use its amplitude $|y(t)|$ as the data to drive our activity recognition and analytics. Based on the radar principle, we define two types of windows: *short window* and *long window*. A long window is composed of a number of short windows each of which consists of n consecutive samples of $|y(t_i)|$. The short window size n is empirically determined and it should be small enough to capture significant body part movements in an activity. We denote the data in the short window k as $A_k = (|y_{k0}(t)|, |y_{k1}(t)|, ...|y_{k(n-1)}(t)|)$, and the data of a long window S as $S = (A_0, A_1, ...A_l)$. In particular, our solution has the short window consisting of 500 consecutive measurements of $|y(t)|$. The long window spans an PRI consisting of 16 short windows for 8,000 measurements in total. Therefore, we have the short window data $A_k = (|y_{k0}(t)|, |y_{k1}(t)|, ...|y_{k(499)}(t)|)$, and the long window data $S = (A_0, A_1, ...A_{15})$.

5.2 Human Activity Recognition

With radar signal data collected and formatted into short and long windows, the activity recognition of *HARAR* is perfomed with three core modules: (1) data preprocessing that extracts the signal $\sum_0^{N-1} \rho_i(t) r_i(t) x(t)$ reflected by human body parts from the received signal mixture with noise, or namely filters out the background noise $n(t)$ from $y(t)$, (2) coarse activity classification that categorizes preprocessed signals into two groups:

high-frequency and low-intensity activities, with a lowpass filter algorithm, and (3) fine activity recognition that recognizes specific activities in each group. The recognition procedure is illustrated as in Fig. 6.

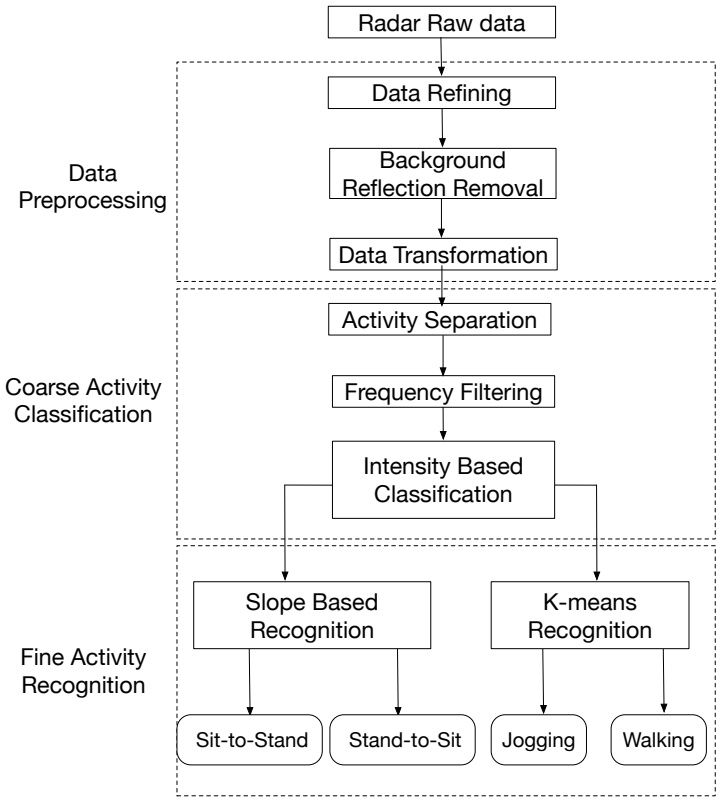

Figure 6: Block diagram of the Activity Recognition system design

5.2.1 Data Preprocessing

After raw measurements of $y(t)|$ are obtained, these data pass through a chain of signal processing blocks that: 1) filter out the background noise from the received mixture measurements, 2) remove exceptional measurements, and 3) transform the data for feature extraction.

Data Refining: The collected radar signal data is noisy because of: (1) the irregular surfaces of static objects in a background environment, (2) unexpected other signals, and (3) the imperfect mechanical capabilities e.g. reliability, stability, accuracy and resolution

of the off-the shelf device.

The first step of the data preprocessing is to remove those exceptional data in each short window. In our platform, the transmitted radar signal has an upper bound amplitude. After the path loss and multi-path fading, the amplitude of the received radar signal $y(t)$ should be less than that of the transmitted signal $x(t)$. Namely this follows: $|y(t)| < |x(t)|$. Therefore, our algorithm removes a measurement $|y_i(t)|$ if $|y_i(t)| \geqslant |x(t)|$. To keep the same number of measurements in each short window, an interpolated measurement is needed. In our solution, the removed measurement is replaced by the mean of its previous and next measurements: $y_i = mean(y_{i-1}, y_{i+1})$.

Background Reflection Removal: Even after the exceptional data are removed, measurements contain both the signal reflected by body parts and by a background environment. One important observation is that, since a background environment is static, there should be no change on its reflection in different radar signal intervals. Namely, the variations of measurements across short or long windows should be only incurred by human body activities. Based on this observation, we design a contrastive divergence algorithm [20] to remove the background reflected signal and keep only reflections by body parts.

As shown in Equation (7), the contrastive divergence algorithm first calculates the divergence of all measurements between every two consecutive short windows, namely $|A_{k+1} - A_k|$, and then identifies and denotes the maximal divergence with its index i as D_i, which indicates the activity results in a signal variation peak at the i-th measurement in a short window.

$$D_i = max(|A_{k+1} - A_k|) \tag{6}$$

$$= max(|A_{k+1}[0] - A_k[0])|, ..., |A_{k+1}[n-1] - A_k[n-1]|) \tag{7}$$

Data Transformation: After the activity is captured with the maximal divergence D_i by the contrastive divergence algorithm, we only know its occurrence at a certain measurement moment in a particular short window. To recognize various activities, it is necessary to gain the knowledge of *how the location (namely the occurrence moment) of the maximal divergence D temporally changes across a sequence of short windows.* It is the pattern of the temporal location change of D that indicates various activities. We will use this information across the short windows in each long window S for activity recognition.

It is difficult to obtain the temporal location change of D. We rather transform the temporal location change of D across short windows to a relative *spatial* location change L, which, based on wireless propagation, is defined as:

$$L_i = |I_i - I_0| \times c/2 \tag{8}$$

where I_i and I_0 respectively represent the occurrence indice of D in the i-th and the first short window in a long window, c is the speed of light, and the division by 2 is due to the radar signal round-trip propagation.

The transformation outcome is a vector M of a long window S, which records maximal divergence D and its relative spatial locations across a sequence of short windows.

Suppose each long window S contains n short windows. M has a format as:

$$M = \begin{bmatrix} (D_0, L_0) & (D_1, L_1) & \cdots & (D_{n-1}, L_{n-1}) \end{bmatrix} \tag{9}$$

5.2.2 Coarse Activity Classification

With the preprocessed data M across long windows, we further process the data to extract the high-level activity features, which includes (1) identifying individual activity occurrences, and (2) extracting body activity frequencies for each activity.

Activity Separation: With a period of time T modulated into a number of long windows $(S_0, S_1...)$, the activities occurring over the time period T can be separated with the data $(M_0, M_1...)$ over the long windows $(S_0, S_1...)$. Observing that between two activities there is only static "silent" background environment that results in 0 for the maximal divergence D, the activities are thus separated by a streak of "0"s in M. We define the activity length as a time-span T_{span} of the activity:

$$T_{span} = t_{end} - t_{start} \tag{10}$$

where t_{end} denotes the time when the activity ends, and t_{start} denotes the time when the activity starts. For instance, if an M has the data of (0, 0, D_i, D_{i+1}, D_{i+2}, $\cdots$, D_n, 0, 0), the time of D_i is counted as t_{start}, and the time of D_n is t_{end}. The time-span T_{span} is used to indicate a single complete activity phase. Then, deep features will be extracted over each T_{span} for the activity recognition.

Frequency Filtering: From frequency domain perspectives, human activities normally occur at low frequencies. For example, walking or running at certain velocities that

cannot be as fast as a car, standing or sitting falling into a motion speed and acceleration range that are relatively small [23]. To further improve the recognition accuracy, we use a lowpass filter, Butterworth filter (BWF) algorithm [3], to remove all unwanted components from the preprocessed data and keep only the human activity data for the recognition. BWF filter focuses on eliminating noises and keeping fundamental activity motion information. It works as:

$$H(\omega)^2 = \frac{{G_0}^2}{1 + (\frac{\omega}{\omega_c})^{2n}} \tag{11}$$

where ω_c denotes cutoff frequency, ω denotes the input frequency. G_0 denotes the DC gain (the gain at zero frequency), which is a constant, and n represents the order of filter. The output $H(\omega)^2$ denotes the gain of the BWF working on the signal of frequency ω. After the cutoff frequency ω_c is set, the processed signal data contains only those frequencies less than or equal to ω_c.

In our solution, the vectors M of each separated activity are the input data to the BWF filter. The cutoff frequency ω_c is determined according to the research outcomes on frequency and velocity of people walking [23].

Intensity Based Classification: In activities: *sit-to-stand, stand-to-sit, walking,* and *jogging*, *sit-to-stand* and *stand-to-sit* obviously do not result in as much intensity as *walking* and *jogging* do. Therefore, they expect to incur much smaller relative spatial location changes L than those of *walking* and *jogging*. Therefore, the output of our BWF algorithm is classified into two categories: *low-intensity activities* and *high-intensity activities* based on their relative spatial location changes L. From our extensive tests, L=1.0 is the best threshold to differentiate these two categories. The fine activity recognition

within these two categories is then performed by exploiting deep features as described next.

5.3 Fine Activity Recognition

After activities into two categories upon the motion intensity in the data, the final activity recognition is performed within each class.

5.3.1 Low-Intensity Activity Recognition

The low-intensity activities includes *sit-to-stand* and *stand-to-sit*. From the numerous measurements in these two activities, we have observed that *they result in two opposite slopes in the relative spatial location changes* L as defined in Equation (8): *ascending* and *descending*. This is because, no matter how the radar sensor is deployed, these two activities result in either the body approaching to the sensor or departing from the sensor in the space. In our system where the sensor is deployed on a desk that faces to the upper body portion, *sit-to-stand* generates an approaching style leading to a descending slope in L while *stand-to-sit* incurs an ascending slope.

The slope detection is performed as follows. These activities are first separated and extracted over a sequence of long windows with the time spanning algorithm in Equation 10. As a result, each T_{span} contains an activity of either *sit-to-stand* or *stand-to-sit*. Since the slope of each activity can be only either ascending or descending, the slope is determined according to the relative spatial location changes L in the T_{span}.

5.3.2 High-Intensity Activity Recognition

For high-intensity activities *walking* and *jogging*, they obviously differ in motion velocity. Thus, we use the velocity as the key feature to differentiate them. Reasonably assuming the velocity of the body does not change widely within a few of long windows, with the observation that the motion results in contrastive divergence D occurring at various indices in a sequence of long windows, the velocity V_i in the i-th long window of an activity follows:

$$V_i \propto \frac{1}{|I_i - I_{i-1}|} \tag{12}$$

where I_i is the maximal contrastive divergence occurring index in the long window. Therefore, rather than getting the actual velocity, we use $|I_i - I_{i-1}|$ as a "relative velocity" to represent the actual velocity in recognition. After we obtain the velocity, we then use k-means classification algorithm to different the *walking* and *jogging* activities.

6 PERFORMANCE EVALUATION

In this section, we have extensively evaluated the performance for two designed motion detection algorithms in a real environment.

6.1 Localization Evaluation

We have extensively evaluated the performance of DiLT in a real environment.

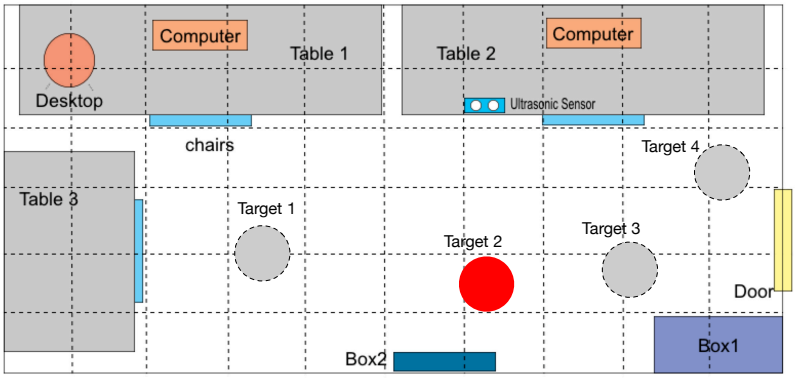

Figure 7: Experiment Environment

6.1.1 Experiment Settings

The evaluation has been performed in a research laboratory (Robert Bell Hall Room 480). This room is complex and noisy, including large furnitures such as desks, tables, chairs and desktops, as well as small devices such as mouses, cables and bottles. A diagram of the experiment setup is illustrated in Figure 7. Small devices are not plotted on the figure. The designed sensor system is placed on the desk 2. In experiments, four locations marked on the figure are tested.

6.1.2 Case#1: System Validation and Background Representation

Our first evaluation scenario is to validate the effectiveness of DiLT in using the distance to represent the static environment background. In the experiment, the room environment has been scanned for ten loops by DiLT. No human or other mobile object exists in the room. With the sensor is incrementally rotated by 2° each time. One scan loop from 0° to 180° results in 91 pieces of data collected. The collected raw distance data are plotted in Figure 8(a). Each line presents one scan loop in these ten loops. As we can observe from the figure, the collected raw data contains a significant amount of noise

that is indicated by the fluctuations along these lines. After the raw data are cleaned by the data processing algorithm of DiLT described in Section 4.3.1, the environment data is finally shown clean as in Figure 8(b). This test indicates that *the data processing to remove exceptional data is effective in using distance to represent the environment.*

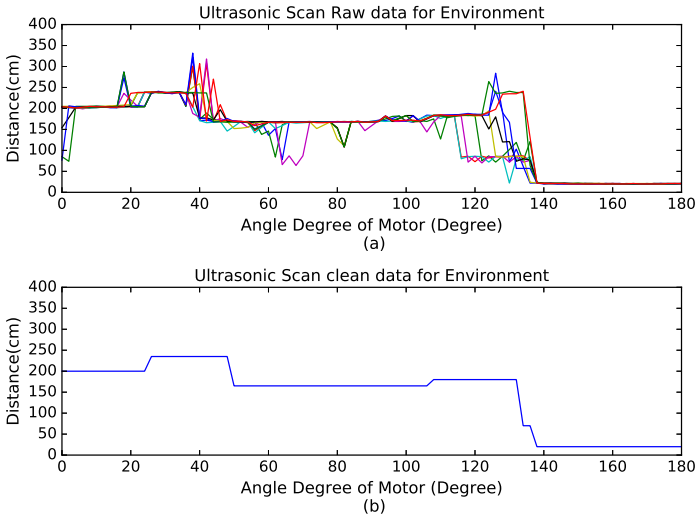

Figure 8: Static Environment Background Representation

6.1.3 Case#2: User Localization

The second case evaluates the localization effectiveness of DiLT. After the static background data is processed in the case#1, DiLT starts a new scanning task to detect a human object in the room. In this test, the human object stays across four different locations. When the user is in each location, DiLT scans for ten loops as it did for the static background. The collected raw sensor data of these four targets are first processed to remove exceptional data, and then analyzed by the localization algorithm as in Section 4.4.1 . Figure 9 plots the localization results. The x-y plane shows the location information in term of (distance, angle), and z-axis is the divergence between the data with user and the static background data without the user. As we can observe, there

are four obvious divergent parts that indicate the four predicted user locations (distance, angle) corresponding to the marked locations in Figure 7.

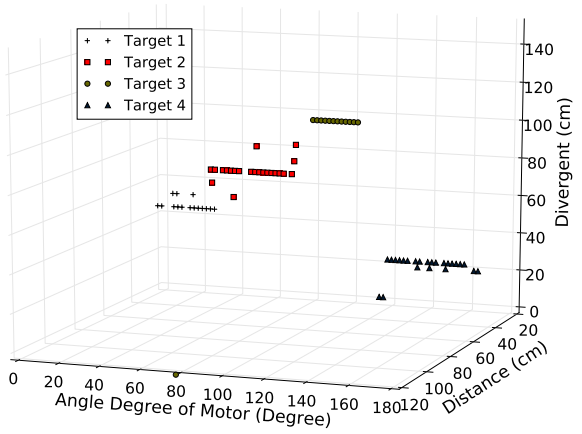

Figure 9: Localization for Different Positions

6.1.4 Case#3: Location Tracking

This test case is to evaluate the capability of DiLT in tracking the change of the user location. As discussed in Section 4.4.2, the location tracking is performed upon detection of the user location in a serial of successive scan tasks by focusing on the adjacent region. We select the target#2 in Figure 7 as an example to test the tracking algorithm. We collect the data that has the user involved in four successive scan tasks t_1, t_2, t_3, and t_4. The predicted locations in these four tasks are shown in Figure 10.

6.1.5 Case#4: Accuracy

In addition, we have evaluated the prediction accuracy of DiLT in localization. The accuracy indicates how close a predicted target location to its true location in the test

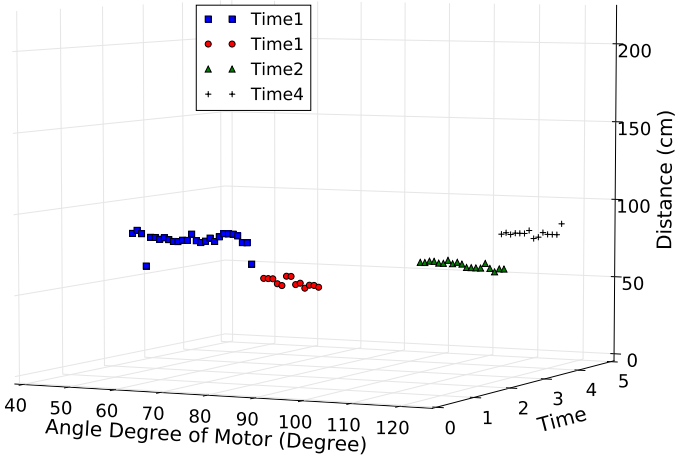

Figure 10: Location Tracking

room. We have performed the test for 20 times with user at various locations. Table 2 shows the results of the accuracy and deviation of localization by DiLT.

Table 1: Localization Accuracy

Parameter	Value
Accuracy in percentage	76%
Max deviation	±30 (degree)
Min deviation	±5 (degree)
Mean deviation	±12 (degree)

6.2 Activity Recognition Evaluation

We have extensively evaluated the performance of our *HARAR* in a real environment.

6.2.1 Experiment Setting

The experiments have been performed in the Intelligent Computing and Communication Systems research lab RB305 at Ball State University. The lab room has a "L" shape with 10 desks and some other furnitures such as chairs and cabinets. This room has a wide space to allow people to perform test activities as in regular life. Figure 11 shows the

deployment of the test platform in a part of the room. In *sit-to-stand* and *stand-to-sit*, the human object is 2m away from the radar sensor. In *walking* and *jogging*, the human object moves up to 5m from the radar sensor.

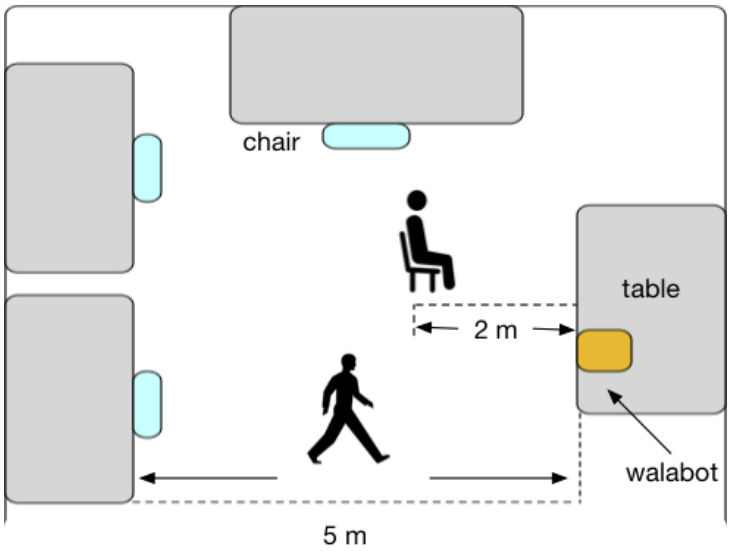

Figure 11: Experiment Setting

6.2.2 Activity Detection and Patterns

The first evaluation is to verify the effectiveness of *HARAR* in detecting human body motions and forming activity patterns. In the experiments, the human objects have performed each of those four activities for 30 seconds. During *walking* and *jogging*, the human objects move back and forth. To magnify the details, we have selected the data of the first 150 out of about 430 long windows in each activity. The results are plotted in Figure 12, where y-axis of all sub-figures represents the received signal power level in the scale of the radar, and x-axis tells the long window number. It can be observed that (1) *sit-to-stand* and *stand-to-sit* both result in clearly lower frequencies and received power levels than *walking* and *jogging*, and (2) the data of *walking* and *jogging* seems continuous while *sit-to-stand* and *stand-to-sit* are bursty.

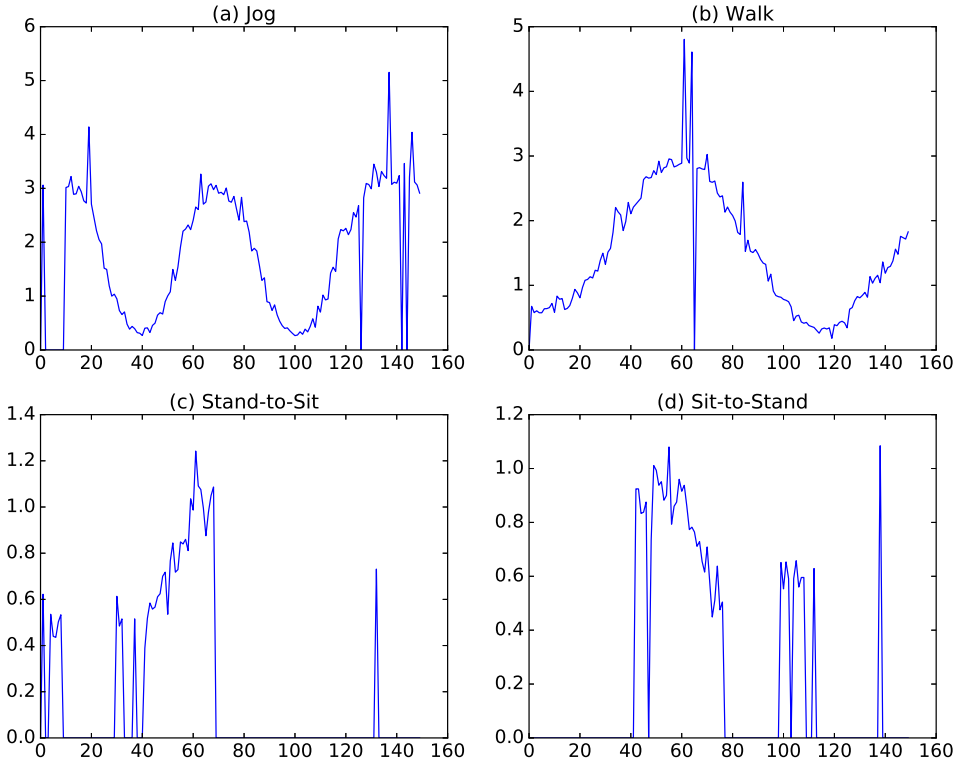

Figure 12: Activity Detection and Patterns

6.2.3 Coarse Activity Classification

This experiment evaluates the effectiveness of using the BWF frequency filter algorithm and the motion intensity to classify the activities into two coarse categories. The BWF filter algorithm has been applied to the entire 430 preprocessed activity data in each activity that are collected in the experiments in Section 6.2.2. The cutoff frequency ω_c is loosely set to 3.3 Hz to allow the activity frequencies are captured safely. The results of four activities are illustrated in Figure 13 where y-axis still refers to the relative spatial location change L. From the frequency domain analysis, *sit-to-stand* and *stand-to-sit* have peak frequencies smaller than 3 Hz while those of *walking* and *jogging* are larger than 3 Hz. Meanwhile, *walking* and *jogging* have mostly resulted large relative spatial location changes (0.5, 3) on the y-axis while *sit-to-stand* and *stand-to-sit* have the values smaller than 1.0. This meets the expectation and analysis in Section 5.2.2.

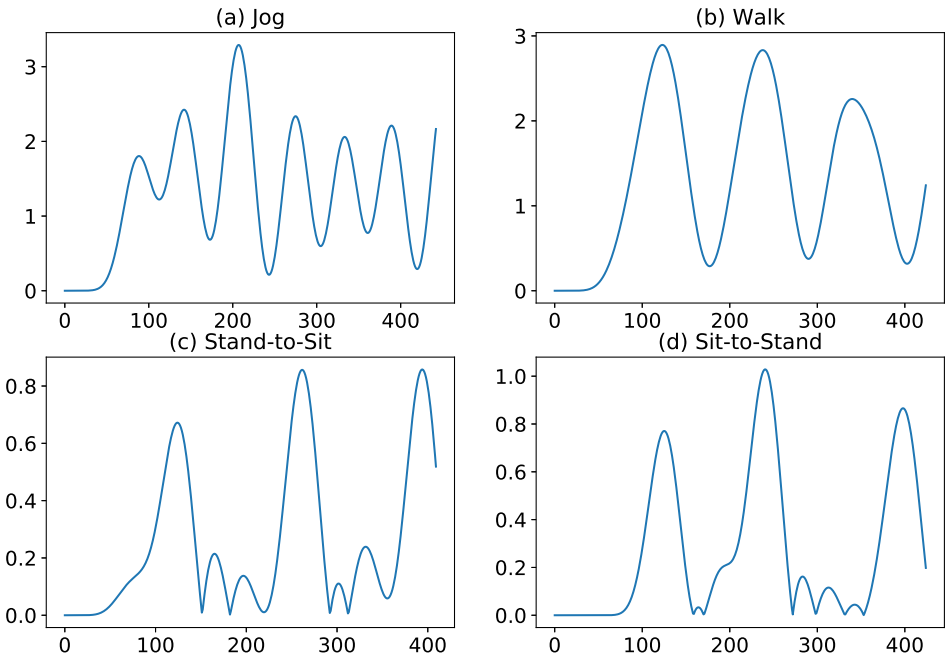

Figure 13: BWF Frequency Filtering on Activity Data

6.2.4 Fine Activity Recognition

We have then performed experiments to evaluate the effectiveness of fine activity recognition algorithms in each coarse activity group.

Low-Frequency Activities In this experiment, the human objects have continuously performed seven *sit-to-stand* and six *stand-to-sit* activities. The data has been first preprocessed to generate the relative spatial location changes L over long windows, which is plotted on the left in Figure 14. We have then used the time spanning algorithm as in Section 5.2.2 to separate activities, and the slope detection algorithm in Section 5.3.1 to determine the activity is *sit-to-stand* or *stand-to-sit*. As a result, *HARAR* can accurately recognize those seven *sit-to-stand* and six *stand-to-sit* activities.

High-Frequency Activities In this experiment, the human objects have walked back and forth at random speeds in one minute and then jogged for another minute. After preprocessing, we have calculated the relative velocities as in Section 5.3.2, which is

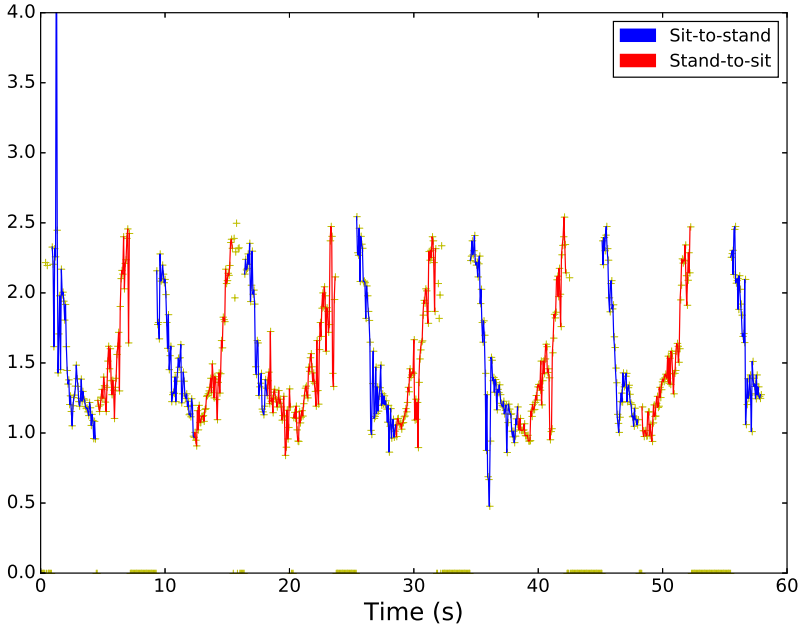

Figure 14: Relative spatial locations of low-frequency activities

plotted in Figure 15 with 880 data points where the dots in blue are from *jogging* and red dots are from *walking*. We can observe that the relative velocity clearly shows advantages in differing these two activities.

6.2.5 Accuracy

Finally, we have evaluated the prediction accuracy of *HARAR* in activity recognition. We have performed all four types of activities totally for 80 tests: 23 walkings, 20 joggings, 17 sit-to-stands and 20 stand-to-sits. Each activity has lasted for 60 seconds. All the data have been mixed together and preprocessed. Then they are passed in the frequency filtering and classified into two coarse categories. Then the final recognition has been performed in each category. The recognition between *walking* and *jogging* is performed with k-means algorithm based on the relative velocity as in Section 5.3.2. The results are shown in Table 2. The accuracies for *walking, jogging, sit-to-stand*, and *stand-to-sit* are respectively: 82.6%, 95%, 82.3%, and 80%. The overall accuracy is 85%. The

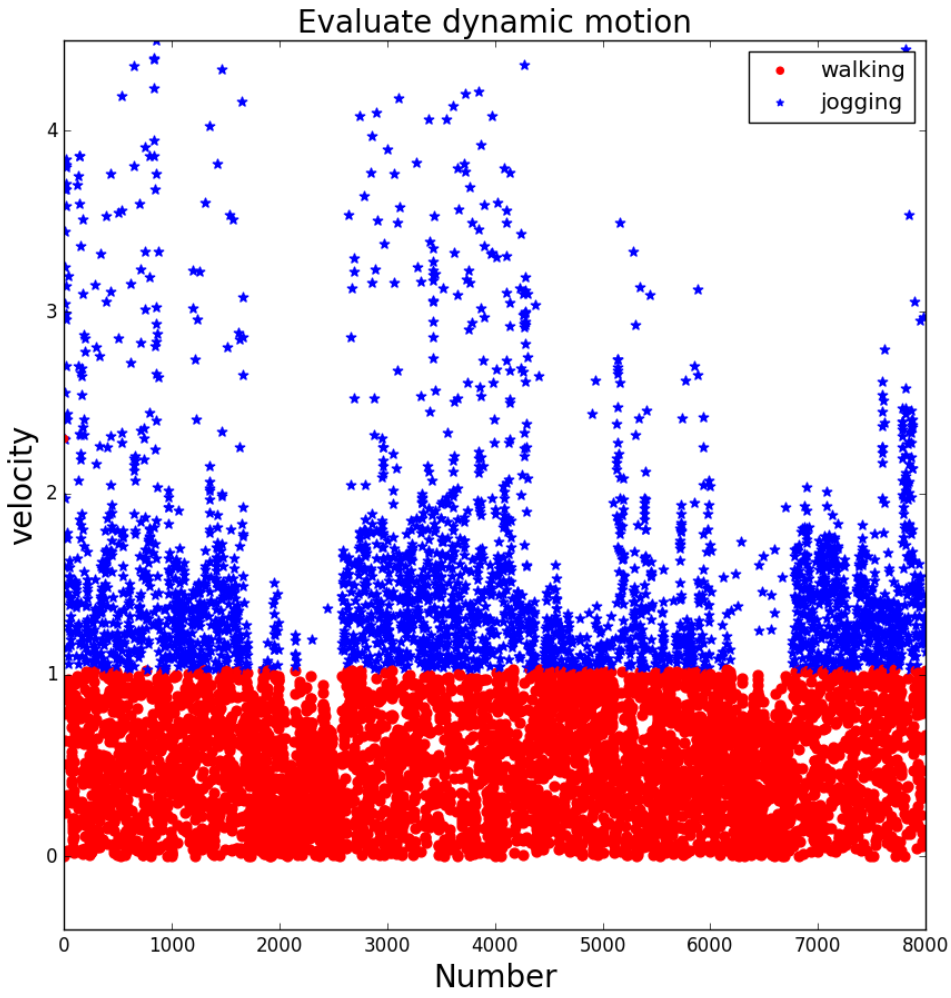

Figure 15: Relative velocity of high-frequency activities

performance seems not very high, but it is outstanding with only the radar sensor without extra equipment.

Table 2: Prediction Accuracy

Recognized/Actual	walking	jogging	sit-to-stand	stand-to-sit
walking	19	1	0	0
jogging	1	19	0	0
sit-to-stand	2	0	14	4
stand-to-sit	1	0	3	16
Accuracy	82.6%	95%	82.3%	80%

7 CONCLUSION

In this work, we describe two novel approaches to track users location system, DiLT, and recognize the human activates method *HARAR* through commodity off-the-shelf radar sensors. Through the mechanical beamforming supported by a rotational ultrasonic scanning, DiLT preprocesses and analyzes collected data to determine the location of a user and track any location changes. *HARAR* measures the human activities with radar signals at very high frequency sampling to capture activity dynamics. Window-based signal processing algorithms are designed to remove background environment noise.

For the localization systerm, DiLT collects data by using commodity off-the-shelf ultrasonic sensors for minimal invasion of privacy, while adopting signal processing techniques to reveal detail dynamics embedded in the sensed data. DiLT consists of a mechanical ultrasonic beam-forming design for omni-space sensing, a contrastive divergence learning to localize a user and a binary back-off algorithm to track the motion of the user.

For the recognition system, this solution designs a set of data preprocessing algorithms, including a data refining algorithm to filter outlier data, a contrastive divergence algorithm to remove background static reflection, and a transformation algorithm to convert the signal data into feature-rich spatial location changes. This solution also develops schemes to separate a collection of various activities into individuals. A low-pass frequency filter is designed to remove unwanted noisy data and the motion intensity is used to classify the activities into two high-level groups. It uses a slope-based approach and a k-means clustering to further finely recognize each activity. From many experiments monitoring laboratory room activity, our proposed recognition algorithms have demonstrated high effectiveness and accuracy in the real environment.

Considering the future work, we expect higher accuracy by using our mode. To improve the accuracy, we can increase the input dataset size and it is necessary to use more sensor data to learn user gestures. We will utilize more machine learning algorithm even the deep learning to test our experiments. Moreover, we will continue explore the depth motion features for recognizing more specific motions in daily lives. Through different kinds of activities, we can learn the users motion feature better. Also, based on our research, the algorithm for user identify will add in the next stage system design, in order to provide more convenient lives in smart homes.

REFERENCES

[1] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 837–846. ACM, 2015.

[2] E. Akinlabi, M. Shukla, S. Akinlabi, S. Kanyanga, and C. Chizyuka. Forming behaviour of steel sheets after mechanical and laser beam forming. *Lasers in Engineering (Old City Publishing)*, 29, 2014.

[3] A. S. Ali, A. G. Radwan, and A. M. Soliman. Fractional order butterworth filter: active and passive realizations. *IEEE Journal on emerging and selected topics in circuits and systems*, 3(3):346–354, 2013.

[4] I. Amir and K. Annamalai. Methods and systems for synchronized ultrasonic real time location, Dec. 10 2013. US Patent 8,604,909.

[5] D. Chakraborty, K. Sharma, R. K. Roy, H. Singh, and T. Bezboruah. Android application based monitoring and controlling of movement of a remotely controlled robotic car mounted with various sensors via bluetooth. In *Advances in Electrical, Electronic and Systems Engineering (ICAEES), International Conference on*, pages 170–175. IEEE, 2016.

[6] K. Chen, H.-S. Lee, A. P. Chandrakasan, and C. G. Sodini. Ultrasonic imaging transceiver design for cmut: A three-level 30-vpp pulse-shaping pulser with improved efficiency and a noise-optimized receiver. *IEEE Journal of Solid-State Circuits*, 48(11):2734–2745, 2013.

[7] S. Choi, E. Kim, and S. Oh. Human behavior prediction for smart homes using deep learning. In *RO-MAN, 2013 IEEE*, pages 173–179. IEEE, 2013.

[8] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. LaMarca, L. LeGrand, A. Rahimi, A. Rea, G. Bordello, B. Hemingway, et al. The mobile sensing platform: An embedded activity recognition system. *IEEE Pervasive Computing*, 7(2), 2008.

[9] S. Das and D. Cook. Designing smart environments: A paradigm based on learning and prediction. *Pattern Recognition and Machine Intelligence*, pages 80–90, 2005.

[10] A. De Angelis, A. Moschitta, P. Carbone, M. Calderini, S. Neri, R. Borgna, and M. Peppucci. Design and characterization of a portable ultrasonic indoor 3-d positioning system. *IEEE Transactions on Instrumentation and Measurement*, 64(10):2616–2625, 2015.

[11] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer. Monitoring activities of daily living in smart homes: Understanding human behavior. *IEEE Signal Processing Magazine*, 33(2):81–94, 2016.

[12] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook. Simple and complex activity recognition through smart phones. In *Intelligent Environments (IE), 2012 8th International Conference on*, pages 214–221. IEEE, 2012.

[13] A. Dixit and A. Naik. Use of prediction algorithms in smart homes. *International Journal of Machine Learning and Computing*, 4(2):157, 2014.

[14] J. Goto, T. Kidokoro, T. Ogura, and S. Suzuki. Activity recognition system for watching over infant children. In *RO-MAN, 2013 IEEE*, pages 473–477. IEEE, 2013.

[15] J. P. Gupta, P. Dixit, and V. B. Semwal. Analysis of gait pattern to recognize the human activities. *IJIMAI*, 2(7):7–16, 2014.

[16] P. Gupta and T. Dallas. Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Transactions on Biomedical Engineering*, 61(6):1780–1786, 2014.

[17] E.-T. Ha, T.-K. Kim, D.-K. Ahn, S.-H. Jeong, I.-R. Yoon, and S.-H. Han. A stable control of legged robot based on ultrasonic sensor. In *Control, Automation and Systems (ICCAS), 2015 15th International Conference on*, pages 1256–1258. IEEE, 2015.

[18] P. M. Hansen, M. Hemmsen, A. Brandt, J. Rasmussen, T. Lange, P. S. Krohn, L. Lönn, J. A. Jensen, and M. B. Nielsen. Clinical evaluation of synthetic aperture sequential beamforming ultrasound in patients with liver tumors. *Ultrasound in medicine & biology*, 40(12):2805–2810, 2014.

[19] J. A. Hartigan and M. A. Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[20] G. E. Hinton. Training products of experts by minimizing contrastive divergence. *Training*, 14(8), 2006.

[21] F. Ijaz, H. K. Yang, A. W. Ahmad, and C. Lee. Indoor positioning: A review of indoor ultrasonic positioning systems. In *Advanced Communication Technology (ICACT), 2013 15th International Conference on*, pages 1146–1150. IEEE, 2013.

[22] A. Jalal and S. Kamal. Real-time life logging via a depth silhouette-based human activity recognition system for smart home services. In *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on*, pages 74–80. IEEE, 2014.

[23] T. Ji et al. Frequency and velocity of people walking. *Structural Engineer*, 84(3):36–40, 2005.

[24] B. Jokanovic, M. Amin, and F. Ahmad. Radar fall motion detection using deep learning. In *Radar Conference (RadarConf), 2016 IEEE*, pages 1–6. IEEE, 2016.

[25] J.-G. Juang, Y.-C. Yang, and J.-A. Wang. Exploring an unknown environment using ultrasonic and infrared sensors. *Sensors and Materials*, 28(9):991–1004, 2016.

[26] Y. Kim, S. Ha, and J. Kwon. Human detection using doppler radar based on physical characteristics of targets. *IEEE Geoscience and Remote Sensing Letters*, 12(2):289–293, 2015.

[27] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev. Soli: Ubiquitous gesture sensing with millimeter wave radar. *ACM Transactions on Graphics (TOG)*, 35(4):142, 2016.

[28] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher. Activity recognition and monitoring using multiple sensors on different body positions. In *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*, pages 4–pp. IEEE, 2006.

[29] M. S. K. Mishra, F. JTMCOE, and K. Bhagat. A survey on human motion detection and surveillance. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume*, 4, 2015.

[30] S. C. Mukhopadhyay. Wearable sensors for human activity monitoring: A review. *IEEE sensors journal*, 15(3):1321–1330, 2015.

[31] Y. Nishida, H. Aizawa, T. Hori, N. H. Hoffman, T. Kanade, and M. Kakikura. 3d ultrasonic tagging system for observing human activity. In *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on*, volume 1, pages 785–791. IEEE, 2003.

[32] Y. Nishida, T. Hori, S.-i. Murakami, and H. Mizoguchi. Minimally privacy-violative system for locating human by ultrasonic radar embedded on ceiling. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 2, pages 1549–1554. IEEE, 2004.

[33] D. Sloo, N. Webb, M. L. Rogers, A. M. Fadell, J. Lee, S. Le Guen, and A. W. Goldenson. Smart-home hazard detector providing useful follow up communications to detection events, Mar. 24 2015. US Patent 8,988,232.

[34] X. Su, H. Tong, and P. Ji. Activity recognition with smartphone sensors. *Tsinghua Science and Technology*, 19(3):235–249, 2014.

[35] Q. Wan, Y. Li, C. Li, and R. Pal. Gesture recognition for smart home applications using portable radar sensors. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 6414–6417. IEEE, 2014.

[36] M. Yamamoto, N. Ajiki, and T. Nakazawa. Ultrasonic diagnosis arrangements for comparing same time phase images of a periodically moving target, Jan. 26 2016. US Patent 9,241,684.

[37] Y. Yan, E. Ricci, G. Liu, and N. Sebe. Egocentric daily activity recognition via multitask clustering. *IEEE Transactions on Image Processing*, 24(10):2984–2995, 2015.

[38] E. Yavari, H. Jou, V. Lubecke, and O. Boric-Lubecke. Doppler radar sensor for occupancy monitoring. In *Silicon Monolithic Integrated Circuits in RF Systems (SiRF), 2013 IEEE 13th Topical Meeting on*, pages 216–218. IEEE, 2013.

[39] Y. Zhang, C. W. de Silva, D. Su, and Y. Xue. Autonomous robot navigation with self-learning for collision avoidance with randomly moving obstacles. In *Computer Science & Education (ICCSE), 2014 9th International Conference on*, pages 117–122. IEEE, 2014.